



# wwPDB X-ray Structure Validation Summary Report

Feb 14, 2017 – 09:02 pm GMT

PDB ID : 3LTG  
Title : Crystal structure of the Drosophila Epidermal Growth Factor Receptor ectodomain complexed with a low affinity Spitz mutant  
Authors : Alvarado, D.; Klein, D.E.; Lemmon, M.A.  
Deposited on : 2010-02-15  
Resolution : 3.40 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the  symbol.

---

The following versions of software and data (see [references](#) ) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

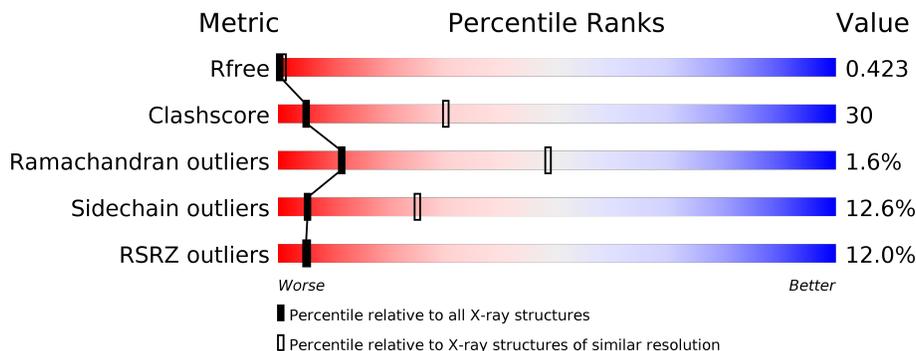
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.40 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	100719	1679 (3.50-3.30)
Clashscore	112137	1832 (3.50-3.30)
Ramachandran outliers	110173	1789 (3.50-3.30)
Sidechain outliers	110143	1789 (3.50-3.30)
RSRZ outliers	101464	1709 (3.50-3.30)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	601	
1	C	601	
2	D	52	

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 8529 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Epidermal growth factor receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	522	4013	2497	709	759	48	0	0	0
1	C	531	4115	2565	732	768	50	0	1	0

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	HIS	-	EXPRESSION TAG	UNP P04412
A	-4	HIS	-	EXPRESSION TAG	UNP P04412
A	-3	HIS	-	EXPRESSION TAG	UNP P04412
A	-2	HIS	-	EXPRESSION TAG	UNP P04412
A	-1	HIS	-	EXPRESSION TAG	UNP P04412
A	0	HIS	-	EXPRESSION TAG	UNP P04412
A	38	GLU	LYS	CONFLICT	UNP P04412
A	230	GLY	ALA	CONFLICT	UNP P04412
A	232	CYS	SER	CONFLICT	UNP P04412
A	359	LEU	ARG	CONFLICT	UNP P04412
A	493	ASN	THR	CONFLICT	UNP P04412
A	590	HIS	-	EXPRESSION TAG	UNP P04412
A	591	HIS	-	EXPRESSION TAG	UNP P04412
A	592	HIS	-	EXPRESSION TAG	UNP P04412
A	593	HIS	-	EXPRESSION TAG	UNP P04412
A	594	HIS	-	EXPRESSION TAG	UNP P04412
A	595	HIS	-	EXPRESSION TAG	UNP P04412
C	-5	HIS	-	EXPRESSION TAG	UNP P04412
C	-4	HIS	-	EXPRESSION TAG	UNP P04412
C	-3	HIS	-	EXPRESSION TAG	UNP P04412
C	-2	HIS	-	EXPRESSION TAG	UNP P04412
C	-1	HIS	-	EXPRESSION TAG	UNP P04412
C	0	HIS	-	EXPRESSION TAG	UNP P04412
C	38	GLU	LYS	CONFLICT	UNP P04412
C	230	GLY	ALA	CONFLICT	UNP P04412

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	232	CYS	SER	CONFLICT	UNP P04412
C	359	LEU	ARG	CONFLICT	UNP P04412
C	493	ASN	THR	CONFLICT	UNP P04412
C	590	HIS	-	EXPRESSION TAG	UNP P04412
C	591	HIS	-	EXPRESSION TAG	UNP P04412
C	592	HIS	-	EXPRESSION TAG	UNP P04412
C	593	HIS	-	EXPRESSION TAG	UNP P04412
C	594	HIS	-	EXPRESSION TAG	UNP P04412
C	595	HIS	-	EXPRESSION TAG	UNP P04412

- Molecule 2 is a protein called Protein spitz.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	D	51	401	259	62	73	7	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	52	ASP	-	EXPRESSION TAG	UNP Q01083





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.75Å 120.15Å 274.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.40 46.58 – 3.40	Depositor EDS
% Data completeness (in resolution range)	82.7 (50.00-3.40) 82.7 (46.58-3.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.79 (at 3.40Å)	Xtrriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.406 , 0.427 0.399 , 0.423	Depositor DCC
$R_{free}$ test set	2886 reflections (11.23%)	DCC
Wilson B-factor (Å <sup>2</sup> )	109.4	Xtrriage
Anisotropy	0.109	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , -9.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.78	EDS
Total number of atoms	8529	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	59.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.88% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality i

### 5.1 Standard geometry i

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.78	9/4103 (0.2%)	0.85	11/5566 (0.2%)
1	C	0.72	3/4211 (0.1%)	0.87	7/5707 (0.1%)
2	D	0.68	0/413	0.77	1/560 (0.2%)
All	All	0.75	12/8727 (0.1%)	0.86	19/11833 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	C	0	1
All	All	0	2

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	490	GLN	C-N	10.65	1.58	1.34
1	A	383	ASN	CG-ND2	-10.62	1.06	1.32
1	A	3	CYS	CB-SG	-10.10	1.65	1.82
1	A	470	ASN	CG-ND2	-10.00	1.07	1.32
1	C	436	ASP	CG-OD1	9.27	1.46	1.25

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	436	ASP	CB-CG-OD2	-17.04	102.97	118.30
1	C	436	ASP	CB-CG-OD1	14.99	131.79	118.30
1	A	267	CYS	CA-CB-SG	-9.31	97.24	114.00
1	C	159	CYS	CA-CB-SG	-9.08	97.66	114.00
2	D	16	CYS	CA-CB-SG	-8.65	98.44	114.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	181	PHE	Mainchain
1	C	158	GLU	Peptide

## 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4013	0	3753	235	2
1	C	4115	0	3906	277	6
2	D	401	0	362	16	0
All	All	8529	0	8021	494	8

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 30.

The worst 5 of 494 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:354:LEU:HD12	1:A:355:ASP:N	1.12	1.44
1:A:354:LEU:CD1	1:A:355:ASP:O	1.76	1.34
1:C:311:HIS:CE1	1:C:351:TYR:CE1	2.21	1.28
1:A:354:LEU:HD12	1:A:355:ASP:CA	1.67	1.25
1:C:311:HIS:CE1	1:C:351:TYR:HE1	1.56	1.23

The worst 5 of 8 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:167:THR:O	1:C:313:GLY:CA[3_554]	1.18	1.02
1:A:48:ASP:CB	1:A:350:ARG:O[4_545]	1.91	0.29
1:C:181:PHE:CD2	1:C:352:ILE:CG2[3_554]	1.99	0.21
1:C:167:THR:O	1:C:313:GLY:C[3_554]	2.02	0.18
1:C:168:HIS:CE1	1:C:352:ILE:N[3_554]	2.03	0.17

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	514/601 (86%)	466 (91%)	38 (7%)	10 (2%)	9	45
1	C	528/601 (88%)	494 (94%)	27 (5%)	7 (1%)	14	53
2	D	49/52 (94%)	46 (94%)	3 (6%)	0	100	100
All	All	1091/1254 (87%)	1006 (92%)	68 (6%)	17 (2%)	11	48

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	383	ASN
1	A	478	CYS
1	C	158	GLU
1	A	212	CYS
1	A	344	ASN

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	440/531 (83%)	394 (90%)	46 (10%)	8	34
1	C	454/531 (86%)	385 (85%)	69 (15%)	3	18
2	D	42/45 (93%)	38 (90%)	4 (10%)	10	38
All	All	936/1107 (85%)	817 (87%)	119 (13%)	5	25

5 of 119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	37	LEU
1	C	127	ASN
1	C	498	ASN
1	C	44	ASN
1	C	92	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	449	GLN
1	C	36	ASN
1	C	493	ASN
1	A	454	GLN
1	A	490	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

### 5.7 Other polymers [i](#)

There are no such residues in this entry.

### 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	522/601 (86%)	0.60	32 (6%) 22 21	35, 60, 83, 106	0
1	C	531/601 (88%)	0.99	93 (17%) 2 2	35, 56, 80, 97	0
2	D	51/52 (98%)	1.04	7 (13%) 3 4	41, 50, 58, 61	0
All	All	1104/1254 (88%)	0.80	132 (11%) 5 5	35, 58, 81, 106	0

The worst 5 of 132 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	153	THR	12.0
1	A	-2	HIS	8.0
1	C	154	ALA	7.2
1	C	506	ASP	5.9
1	C	51	PHE	5.5

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.